(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 January 2002 (03.01.2002)

PCT

(10) International Publication Number
WO 02/00677 A1

(51) International Patent Classification[7]: C07H 21/02, 21/04, C12N 1/20, 15/00, 15/09, 15/63, 15/70, 15/74

(21) International Application Number: PCT/US01/18569

(22) International Filing Date: 7 June 2001 (07.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/209,467 7 June 2000 (07.06.2000) US

(71) Applicant *(for all designated States except US)*: **HUMAN GENOME SCIENCES, INC.** [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **BIRSE, Charles, E.** [GB/US]; 13822 Saddleview Drive, North Potomac, MD 20878 (US). **ROSEN, Craig, A.** [US/US]; 22400 Rolling Hill Lane, Laytonsville, MD 20882 (US).

(74) Agents: **WALES, Michele, M.** et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NUCLEIC ACIDS, PROTEINS, AND ANTIBODIES

(57) Abstract: The present invention relates to novel ovarian related polynucleotides and the polypeptides encoded by these polynucleotides herein collectively known as "ovarian antigens", and the use of such ovarian antigens for detecting disorders of the ovaries and/or breast, particularly the presence of ovarian and/or breast cancer and ovarian and/or breast cancer metastases. More specifically, isolated ovarian associated nucleic acid molecules are provided encoding novel ovarian associated polypeptides. Novel ovarian polypeptides and antibodies that bind to these polypeptides are provided. Also provided are vectors, host cells, and recombinant and synthetic methods for producing human ovarian associated polynucleotides and/or polypeptides. The invention further relates to diagnostic and therapeutic methods useful for diagnosing, treating, preventing and/or prognosing disorders related to the ovaries and/or breast, including ovarian and/or breast cancer, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying agonists and antagonists of polynucleotides and polypeptides of the invention. The present invention further relates to methods and/or compositions for inhibiting the production and function of the polypeptides of the present invention.

WO 02/00677 A1

Nucleic Acids, Proteins, and Antibodies

[0001] This application refers to a "Sequence Listing" that is provided on electronic media in computer readable form pursuant to Administrative Instructions Section 801(a)(i). The Sequence Listing forms a part of this description pursuant to Rule 5.2 and Administrative Instructions Sections 801 to 806, and is hereby incorporated in its entirety.

[0002] The Sequence Listing is provided as an electronic file (PA133PCTSL.txt, 7,347,875 bytes in size, created on 07 June 2001) on four identical compact discs (CD-R), labeled "COPY 1," "COPY 2," "COPY 3," and "CRF." The Sequence Listing complies with Annex C of the Administrative Instructions, and may be viewed, for example, on an IBM-PC machine running the MS-Windows operating system by using the V viewer software, version 2000 (see World Wide Web URL: http://www.fileviewer.com).

Field of the Invention

[0003] The present invention relates to novel ovarian related polynucleotides, the polypeptides encoded by these polynucleotides herein collectively referred to as "ovarian antigens," and antibodies that immunospecifically bind these polypeptides, and the use of such ovarian polynucleotides, antigens, and antibodies for detecting, treating, preventing and/or prognosing disorders of the reproductive system, particularly disorders of the ovaries and/or breast, including, but not limited to, the presence of ovarian and/or breast cancer and ovarian and/or breast cancer metastases. More specifically, isolated ovarian nucleic acid molecules are provided encoding novel ovarian polypeptides. Novel ovarian

nucleic acid molecules are provided encoding novel ovarian polypeptides. Novel ovarian polypeptides and antibodies that bind to these polypeptides are provided. Also provided are vectors, host cells, and recombinant and synthetic methods for producing human ovarian polynucleotides, polypeptides, and/or antibodies. The invention further relates to diagnostic and therapeutic methods useful for diagnosing, treating, preventing and/or prognosing disorders related to the ovaries and/or breast, including ovarian and/or breast cancer, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying agonists and antagonists of polynucleotides and polypeptides of the invention. The invention further relates to methods and/or compositions for inhibiting or promoting the production and/or function of the polypeptides of the invention.

Background of the Invention

[0004] The female reproductive system is comprised of both external and internal organs. The external organs function in permitting sperm to enter the body and protecting the internal genital organs from infection and injury. The internal organs form a pathway (the genital tract) for reproduction, beginning at the ovaries, through the fallopian tubes (oviducts) and uterus, to the birth canal (vagina).

[0005] The sexual and reproductive functions in the female can be divided into two major phases: first, preparation of the body for conception, and second, the gestation and parturition. Gestation and parturition only occur if an ovum becomes fertilized. If fertilization does not occur, the reproductive system undergoes a cycle to ensure frequent readiness for conception and fertilization.

[0006] The complexity of the female reproductive system renders it susceptible to several diseases and disorders. In particular, the ovaries and breast are subject to diseases and/or disorders such as infections, hyperproliferative disorders, as well as regulatory and genetic abnormalities.

Disorders of the Ovary

[0007] A woman's ovaries are located on both sides of the uterus, below the opening of the fallopian tubes (tubes that extend from the uterus to the ovaries). In

addition to producing egg cells for reproduction, the ovaries produce estrogen and progesterone, which affect many of the female characteristics and reproductive functions.

[0008] Anovulation (the absence of egg release by the ovaries) is a serious condition leading to infertility. The exact etiology of anovulation, especially in women with otherwise normal menstrual cycles, is unclear, however several potential causes are under study, including: impaired follicular development (probably due to low or absent estrogen production or binding), normal follicular development with lack of egg release (probably due to progesterone deficiency), or insufficient production of gonadotropin-releasing hormone from the hypothalamus. Current treatments include clomiphene injections or hormonal therapy, although both can lead to serious side effects such as ovarian cancer and ovarian hyperstimulation syndrome.

[0009] Anovulation is also associated with polycyctic ovary syndrome (also known as Stein-Leventhal syndrome). This syndrome is and endocrine disorder characterized by an elevated level of male hormones (androgens). Other than anovulation, symptoms include growth of male-patterned body hair (hirsutism), excessive acne, irregular or absent menses, excessive bleeding, and obesity. Usually, the ovaries appear enlarged and may contain many follicular cysts.

[0010] Ovarian cancer develops most often in women between the ages of 50 and 70. It is the third most common cancer of the female reproductive system, but more women die of ovarian cancers than of any other. Ovaries include a variety of cell types, each of which may give rise to a distinct type of cancer, including, but not limited to, ovarian epithelial cancer, ovarian germ cell tumors, ovarian papillary serous adenocarcinoma, ovarian mucinous adenocarcinoma, ovarian Krukenberg tumor, malignant mixed Mullerian tumors, and ovarian low malignant tumors.

[0011] Other disorders of the ovaries also include, but are not limited to, inflammatory disorders, such as oophoritis (e.g., caused by viral or bacterial infection), ovarian cysts, and autoimmune disorders (e.g., premature ovarian failure and autoimmune oophoritis).

Disorders of the Breast

[0012] The breast is comprised of different structures, each with its own specific function. One-third of the breast is comprised of fatty tissue. The other two-thirds is made

up of structural components called ducts and lobules. Milk is produced in the lobules and funneled through the ducts to the nipple. Disorders of the breast typically involve the formation of lesions within breast tissue. While many of these lesions are benign in nature, they may lead to cancer if left untreated.

[0013] Benign breast lesions include, for example, cysts, which are non-cancerous, fluid-filled sacs that forma mass within breast tissue. The cause of breast cysts is unknown, though injury may be involved, and their main symptom is pain. While considered harmless, a professional should drain cysts and the fluid examined because cancer of the cyst wall, although quite rare, is possible.

[0014] Other benign breast lesions include fibrous breast lumps (fibroadenomas), breast infection (mastitis), intraductal papilloma, and abscesses. Fibrous breast lumps are small, solid lumps of glandular tissue. These lumps usually appear in young women, often in teenagers, and are easy to remove. Intraductal papilloma are small lumps located within a milk duct, often causing inappropriate discharge from the nipple. Breast abscesses are collections of pus in breast tissue that develop from breast infections that go untreated.

[0015] Breast cancer is the most common cancer among women, other than skin cancer and is the second leading cause of cancer death in women, after lung cancer. The American Cancer Society predicts that there will be about 182,800 new cases of invasive breast cancer in the year 2000 among women in this country and about 40,800 deaths from the disease. Breast cancer also occurs among men, although much less often. It is generally believed that this malignancy arises from a multi step process involving mutations in a relatively small number of genes, perhaps 10 or less. These mutations result in significant changes in the growth and differentiation of breast tissue that allow it to grow independent of normal cellular controls, to metastasize, and to escape immune surveillance. The genetic heterogeneity of most breast cancers suggests that they arise by a variety of initiating events and that the characteristics of individual cancers are due to the collective pattern of genetic changes that accumulate.

[0016] The discovery of new human ovarian associated polynucleotides, the polypeptides encoded by them, and antibodies that immunospecifically bind these polypeptides, satisfies a need in the art by providing new compositions which are useful in

the diagnosis, treatment, prevention and/or prognosis of disorders of the ovaries and/or breast, including, but not limited to, neoplastic disorders (e.g., ovarian Krukenberg tumor, malignant mixed Mullerian tumors, and/or as described under "Hyperproliferative Disorders" below), infectious diseases (e.g., mastitis, oophoritis, and/or as described under "Infectious Diseases" below), and inflammatory diseases (e.g., abcesses and/or as described under "Immune Disorders" below) and as described in "Reproductive System Disorders" below.

Summary of the Invention

[0017] The present invention relates to novel ovarian related polynucleotides, the polypeptides encoded by these polynucleotides herein collectively referred to as "ovarian antigens," and antibodies that immunospecifically bind these polypeptides, and the use of such ovarian polynucleotides, antigens, and antibodies for detecting, treating, preventing and/or prognosing disorders of the reproductive system, particularly disorders of the ovaries and/or breast, including, but not limited to, the presence of ovarian and/or breast cancer and ovarian and/or breast cancer metastases. More specifically, isolated ovarian nucleic acid molecules are provided encoding novel ovarian polypeptides. Novel ovarian polypeptides and antibodies that bind to these polypeptides are provided. Also provided are vectors, host cells, and recombinant and synthetic methods for producing human ovarian polynucleotides, polypeptides, and/or antibodies. The invention further relates to diagnostic and therapeutic methods useful for diagnosing, treating, preventing and/or prognosing disorders related to the ovaries and/or breast, including ovarian and/or breast cancer, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying agonists and antagonists of polynucleotides and polypeptides of the invention. The invention further relates to methods and/or compositions for inhibiting or promoting the production and/or function of the polypeptides of the invention.

Detailed Description

Tables

[0018] Table 1 summarizes some of the polynucleotides encompassed by the invention (including cDNA clones related to the sequences (Clone ID NO:Z), contig sequences (contig identifier (Contig ID:) and contig nucleotide sequence identifier (SEQ ID NO:X)) and further summarizes certain characteristics of these polynucleotides and the polypeptides encoded thereby. The first column provides a unique clone identifier, "Clone ID NO:Z", for a cDNA plasmid related to each ovarian associated contig sequence disclosed in Table 1. The second column provides a unique contig identifier, "Contig ID:" for each of the contig sequences disclosed in Table 1. The third column provides the sequence identifier, "SEQ ID NO:X", for each of the contig polynucleotide sequences disclosed in Table 1. The fourth column, "ORF (From-To)", provides the location (i.e., nucleotide position numbers) within the polynucleotide sequence of SEQ ID NO:X that delineate the preferred open reading frame (ORF) shown in the sequence listing and referenced in Table 1 as SEQ ID NO:Y (column 5). Column 6 lists residues comprising predicted epitopes contained in the polypeptides encoded by each of the preferred ORFs (SEQ ID NO:Y). Identification of potential immunogenic regions was performed according to the method of Jameson and Wolf (CABIOS, 4:181-186 (1988)); specifically, the Genetics Computer Group (GCG) implementation of this algorithm, embodied in the program PEPTIDESTRUCTURE (Wisconsin Package v10.0, Genetics Computer Group (GCG), Madison, Wisc.). This method returns a measure of the probability that a given residue is found on the surface of the protein. Regions where the antigenic index score is greater than 0.9 over at least 6 amino acids are indicated in Table 1 as "Predicted Epitopes." In particular embodiments, ovarian associated polypeptides of the invention comprise, or alternatively consist of, one, two, three, four, five or more of the predicted epitopes described in Table 1. It will be appreciated that depending on the analytical criteria used to predict antigenic determinants, the exact address of the determinant may vary slightly. Column 7, "Tissue Distribution" shows the expression profile of tissue, cells, and/or cell line libraries which express the polynucleotides of the invention. The first number in column 7 (preceding the colon), represents the tissue/cell source identifier

code corresponding to the code and description provided in Table 4. Expression of these polynucleotides was not observed in the other tissues and/or cell libraries tested. For those identifier codes in which the first two letters are not "AR", the second number in column 7 (following the colon), represents the number of times a sequence corresponding to the reference polynucleotide sequence (e.g., SEQ ID NO:X) was identified in the tissue/cell source. Those tissue/cell source identifier codes in which the first two letters are "AR" designate information generated using DNA array technology. Utilizing this technology, cDNAs were amplified by PCR and then transferred, in duplicate, onto the array. Gene expression was assayed through hybridization of first strand cDNA probes to the DNA array. cDNA probes were generated from total RNA extracted from a variety of different tissues and cell lines. Probe synthesis was performed in the presence of ^{33}P dCTP, using oligo(dT) to prime reverse transcription. After hybridization, high stringency washing conditions were employed to remove non-specific hybrids from the array. The remaining signal, emanating from each gene target, was measured using a Phosphorimager. Gene expression was reported as Phosphor Stimulating Luminescence (PSL) which reflects the level of phosphor signal generated from the probe hybridized to each of the gene targets represented on the array. A local background signal subtraction was performed before the total signal generated from each array was used to normalize gene expression between the different hybridizations. The value presented after "[array code]:" represents the mean of the duplicate values, following background subtraction and probe normalization. One of skill in the art could routinely use this information to identify normal and/or diseased tissue(s) which show a predominant expression pattern of the corresponding polynucleotide of the invention or to identify polynucleotides which show predominant and/or specific tissue and/or cell expression. Column 8, "Cytologic Band," provides the chromosomal location of polynucleotides corresponding to SEQ ID NO:X. Chromosomal location was determined by finding exact matches to EST and cDNA sequences contained in the NCBI (National Center for Biotechnology Information) UniGene database. Given a presumptive chromosomal location, disease locus association was determined by comparison with the Morbid Map, derived from Online Mendelian Inheritance in Man (Online Mendelian Inheritance in Man, OMIM™. McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University (Baltimore, MD) and National Center for Biotechnology Information, National Library of Medicine (Bethesda, MD) 2000. World

Wide Web URL: http://www.ncbi.nlm.nih.gov/omim/). If the putative chromosomal location of the Query overlapped with the chromosomal location of a Morbid Map entry, an OMIM identification number is provided in column 9 labeled "OMIM Disease Reference(s)". A key to the OMIM reference identification numbers is provided in Table 5.

[0019] Table 2 summarizes homology and features of some of the polypeptides of the invention. The first column provides a unique clone identifier, "Clone ID NO:Z", corresponding to a cDNA disclosed in Table 1. The second column provides the unique contig identifier, "Contig ID:" corresponding to contigs in Table 1 and allowing for correlation with the information in Table 1. The third column provides the sequence identifier, "SEQ ID NO:X", for the contig polynucleotide sequences. The fourth column provides the analysis method by which the homology/identity disclosed in the row was determined. Comparisons were made between polypeptides encoded by the polynucleotides of the invention and either a non-redundant protein database (herein referred to as "NR"), or a database of protein families (herein referred to as "PFAM") as further described below. The fifth column provides a description of PFAM/NR hits having significant matches to a polypeptide of the invention. Column six provides the accession number of the PFAM/NR hit disclosed in the fifth column. Column seven, "Score/Percent Identity", provides a quality score or the percent identity, of the hit disclosed in column five. Columns 8 and 9, "NT From" and "NT To" respectively, delineate the polynucleotides in "SEQ ID NO:X" that encode a polypeptide having a significant match to the PFAM/NR database as disclosed in the fifth column. In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence encoded by the polynucleotides in SEQ ID NO:X as delineated in columns 8 and 9, or fragments or variants thereof.

[0020] Table 3 provides polynucleotide sequences that may be disclaimed according to certain embodiments of the invention. The first column provides a unique clone identifier, "Clone ID NO:Z", for a cDNA clone related to ovarian associated contig sequences disclosed in Table 1. The second column provides the sequence identifier, "SEQ ID NO:X", for contig polynucleotide sequences disclosed in Table 1. The third column provides the unique contig identifier, "Contig ID", for contigs disclosed in Table 1. The fourth column provides a unique integer 'a' where 'a' is any integer between 1 and

the final nucleotide minus 15 of SEQ ID NO:X, represented as "Range of a", and the fifth column provides a unique integer 'b' where 'b' is any integer between 15 and the final nucleotide of SEQ ID NO:X, represented as "Range of b", where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:X, and where b is greater than or equal to a + 14. For each of the polynucleotides shown as SEQ ID NO:X, the uniquely defined integers can be substituted into the general formula of a-b, and used to describe polynucleotides which may be preferably excluded from the invention. In certain embodiments, preferably excluded from the polynucleotides of the invention (including polynucleotide fragments and variants as described herein and diagnostic and/or therapeutic uses based on these polynucleotides) are at least one, two, three, four, five, ten, or more of the polynucleotide sequence(s) having the accession number(s) disclosed in the sixth column of this Table. In further embodiments, preferably excluded from the invention are the specific polynucleotide sequence(s) contained in the clones corresponding to at least one, two, three, four, five, ten, or more of the available material having the accession numbers identified in the sixth column of this Table.

[0021] Table 4 provides a key to the tissue/cell source identifier code disclosed in Table 1, column 7. Column 1 provides the key to the tissue/cell source identifier code disclosed in Table 1, Column 7. Columns 2-5 provide a description of the tissue or cell source. Codes corresponding to diseased tissues are indicated in column 6 with the word "disease". The use of the word "disease" in column 6 is non-limiting. The tissue or cell source may be specific (e.g. a neoplasm), or may be disease-associated (e.g., a tissue sample from a normal portion of a diseased organ). Furthermore, tissues and/or cells lacking the "disease" designation may still be derived from sources directly or indirectly involved in a disease state or disorder, and therefore may have a further utility in that disease state or disorder. In numerous cases where the tissue/cell source is a library, column 7 identifies the vector used to generate the library.

[0022] Table 5 provides a key to the OMIM™ reference identification numbers disclosed in Table 1, column 9. OMIM reference identification numbers (Column 1) were derived from Online Mendelian Inheritance in Man (Online Mendelian Inheritance in Man, OMIM™. McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University (Baltimore, MD) and National Center for Biotechnology Information, National Library of Medicine, (Bethesda, MD) 2000. World Wide Web URL:

http://www.ncbi.nlm.nih.gov/omim/). Column 2 provides diseases associated with the cytologic band disclosed in Table 1, column 8, as determined from the Morbid Map database.

[0023] Table 6 summarizes ATCC Deposits, Deposit dates, and ATCC designation numbers of deposits made with the ATCC in connection with the present application.

[0024] Table 7 shows the cDNA libraries sequenced, tissue source description, vector information and ATCC designation numbers relating to these cDNA libraries.

Definitions

[0025] The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

[0026] In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide. The term "isolated" does not refer to genomic or cDNA libraries, whole cell total or mRNA preparations, genomic DNA preparations (including those separated by electrophoresis and transferred onto blots), sheared whole cell genomic DNA preparations or other compositions where the art demonstrates no distinguishing features of the polynucleotide sequences of the present invention.

[0027] As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence encoding SEQ ID NO:Y or a fragment or variant thereof; a nucleic acid sequence contained in SEQ ID NO:X (as described in column 3 of Table 1) or the complement thereof; or a cDNA sequence contained in Clone ID NO:Z (as described in column 1 of Table 1 and contained within a library deposited with the ATCC). For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having an amino acid sequence encoded by a polynucleotide of the invention as broadly defined (obviously excluding poly-

Phenylalanine or poly-Lysine peptide sequences which result from translation of a polyA tail of a sequence corresponding to a cDNA).

[0028] As used herein, an "ovarian antigen" refers collectively to any polynucleotide disclosed herein (e.g., a nucleic acid sequence contained in SEQ ID NO:X or the complement therof, or cDNA sequence contained in Clone ID NO:Z, (e.g., a nucleic acid sequence contained in SEQ ID NO:X or the complement thereof, or cDNA sequence contained in Clone ID NO:Z, and fragments or variants thereof as described herein) or any polypeptide disclosed herein (e.g., an amino acid sequence contained in SEQ ID NO:Y, an amino acid sequence encoded by SEQ ID NO:X, or the complement thereof, an amino acid sequence encoded by the cDNA sequence contained in Clone ID NO:Z, and fragments or variants thereof as described herein). These ovarian antigens have been determined to be predominantly expressed in ovarian tissues, including normal or diseased tissues (as shown in Table 1 column 7 and Table 4).

[0029] In the present invention, "SEQ ID NO:X" was often generated by overlapping sequences contained in multiple clones (contig analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X is deposited at Human Genome Sciences, Inc. (HGS) in a catalogued and archived library. As shown, for example, in column 1 of Table 1, each clone is identified by a cDNA Clone ID (identifier generally referred to herein as Clone ID NO:Z). Each Clone ID is unique to an individual clone and the Clone ID is all the information needed to retrieve a given clone from the HGS library. Furthermore, certain clones disclosed in this application have been deposited with the ATCC on June 5, 2000 and were given ATCC Deposit Nos. PTA-1982 and PTA-1985. In addition to the individual cDNA clone deposits, most of the cDNA libraries from which the clones were derived were deposited at the American Type Culture Collection (hereinafter "ATCC"). Table 7 provides a list of the deposited cDNA libraries. One can use the Clone ID NO:Z to determine the library source by reference to Tables 6 and 7. Table 7 lists the deposited cDNA libraries by name and links each library to an ATCC Deposit. Library names contain four characters, for example, "HTWE." The name of a cDNA clone (Clone ID NO:Z) isolated from that library begins with the same four characters, for example "HTWEP07". As mentioned below, Table 1 correlates the Clone ID NO:Z names with SEQ ID NO:X. Thus, starting with an SEQ ID NO:X, one can use Tables 1A, 6 and 7 to determine the corresponding Clone ID NO:Z, which library it came

from and which ATCC deposit the library is contained in. Furthermore, it is possible to retrieve a given cDNA clone from the source library by techniques known in the art and described elsewhere herein. The ATCC is located at 10801 University Boulevard, Manassas, Virginia 20110-2209, USA. The ATCC deposits were made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for the purposes of patent procedure.

[0030] In specific embodiments, the polynucleotides of the invention are at least 15, at least 30, at least 50, at least 100, at least 125, at least 500, or at least 1000 continuous nucleotides but are less than or equal to 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb, 7.5 kb, 5 kb, 2.5 kb, 2.0 kb, or 1 kb, in length. In a further embodiment, polynucleotides of the invention comprise a portion of the coding sequences, as disclosed herein, but do not comprise all or a portion of any intron. In another embodiment, the polynucleotides comprising coding sequences do not contain coding sequences of a genomic flanking gene (i.e., 5' or 3' to the gene of interest in the genome). In other embodiments, the polynucleotides of the invention do not contain the coding sequence of more than 1000, 500, 250, 100, 50, 25, 20, 15, 10, 5, 4, 3, 2, or 1 genomic flanking gene(s).

[0031] A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, or the complement thereof (e.g., the complement of any one, two, three, four, or more of the polynucleotide fragments described herein), the polynucleotide sequence delineated in columns 8 and 9 of Table 2 or the complement thereof, and/or cDNA sequences contained in Clone ID NO:Z (e.g., the complement of any one, two, three, four, or more of the polynucleotide fragments, ans/or the cDNA clone within the pool of cDNA clones deposited with the ATCC, described herein). "Stringent hybridization conditions" refers to an overnight incubation at 42 degree C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65 degree C.

[0032] Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions.

Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency), salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37 degree C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH_2PO_4; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 ug/ml salmon sperm blocking DNA; followed by washes at 50 degree C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

[0033] Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

[0034] Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone generated using oligo dT as a primer).

[0035] The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons.

"Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

[0036] The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth. Enzymol. 182:626-646 (1990); Rattan et al., Ann. N.Y. Acad. Sci. 663:48-62 (1992).)

[0037] "SEQ ID NO:X" refers to a polynucleotide sequence described, for example, in Tables 1A or 2, while "SEQ ID NO:Y" refers to a polypeptide sequence described in column 5 of Table 1. SEQ ID NO:X is identified by an integer specified in column 3 of Table 1. The polypeptide sequence SEQ ID NO:Y is a translated open reading frame (ORF) encoded by polynucleotide SEQ ID NO:X. "Clone ID NO:Z" refers to a cDNA clone described in column 1 of Table 1.

[0038] "A polypeptide having biological activity" refers to a polypeptide exhibiting activity similar to, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention).

[0039] Table 1 summarizes some of the polynucleotides encompassed by the invention (including contig sequences (SEQ ID NO:X) and clones (Clone ID NO:Z) and further summarizes certain characteristics of these polynucleotides and the polypeptides encoded thereby.

Polynucleotides and Polypeptides